The future of industrial antibiotic production

Medema, M.H.; Alam, M.T.; Breitling, R.; Takano, E.

Published in:
Bioengineered Bugs

DOI:
10.4161/bbug.2.4.16114

Publication date:
2011

Bioengineered Bugs 2:4, 230-233; July/August 2011; © 2010 Landes Bioscience

The future of industrial antibiotic production

From random mutagenesis to synthetic biology

Marnix H. Medema,[1,2,†] Mohammad T. Alam,[2,3,†] Rainer Breitling[2,3] and Eriko Takano[1,*]

[1]Department of Microbial Physiology; [2]Groningen Bioinformatics Centre; Groningen Biomolecular Sciences and Biotechnology Institute; University of Groningen; Groningen, The Netherlands; [3]Institute of Molecular, Cell and Systems Biology; College of Medical, Veterinary and Life Sciences; Joseph Black Building; University of Glasgow; Glasgow, UK

[†]These authors contributed equally to this work.

Natural products derived from the secondary metabolism of microbes constitute a cornerstone of modern medicine. Engineering bugs to produce these products in high quantities is a major challenge for biotechnology, which has usually been tackled by either one of two strategies: iterative random mutagenesis or rational design. Recently, we analyzed the transcriptome of a *Streptomyces clavuligerus* strain optimized for production of the β-lactamase inhibitor clavulanic acid by multiple rounds of mutagenesis and selection, and discovered that the observed changes matched surprisingly well with simple changes that have been introduced into these strains by rational engineering. Here, we discuss how in the new field of synthetic biology, random mutagenesis and rational engineering can be implemented complementarily in ways which may enable one to go beyond the status quo that has now been reached by each method independently.

Key words: antibiotics, synthetic biology, random mutagenesis, Streptomyces, clavulanic acid, natural products, secondary metabolism

Submitted: 04/20/11

Revised: 05/09/11

Accepted: 05/10/11

DOI: 10.4161/bbug.2.4.16114

*Correspondence to: Eriko Takano; Email: e.takano@rug.nl

Addendum to: Medema MH, Alam MT, Heijne WHM, van den Berg MA, Müller U, Trefzer A, et al. Genome-wide gene expression changes in an industrial clavulanic acid overproduction strain of *Streptomyces clavuligerus*. Microb Biotechnol 2011; 4:300–5; PMID: 21342474; DOI: 10.1111/j.1751-7915.2010.00226.x.

In our recent paper in *Microbial Biotechnology*,[1] we compared levels of gene expression between a *Streptomyces clavuligerus* clavulanic acid overproduction mutant (DS48802) and the wild-type. The changes caused by random mutagenesis were strikingly similar to those rationally engineered by two strategies that have recently been employed to increase clavulanic acid production in *S. clavuligerus*: redirection of carbon fluxes towards the key clavulanic acid precursor glyceraldehyde-3-phosphate (G3P)[2] and upregulation of pathway-specific activators.[3,4] More precisely, we found the pathway-specific activator genes *ccaR* and *claR* to be more highly expressed in strain DS48802, and we also observed changes in transcript levels of genes associated with glycolysis and the citric acid cycle, which appear to match the intended effect of the recently engineered *gap1* deletion mutant, leading to a redirection of carbon fluxes towards G3P. A third way in which overproduction of chemicals has been achieved by rational engineering is the genomic duplication or even amplification of the biosynthetic gene cluster. Such changes have recently been successfully implemented for overproduction of platensimycin[5] and nikkomycin.[6] Although a similar amplification has been observed previously in a randomly mutagenized kanamycin overproduction strain,[7] we did not observe any amplification of the clavulanic acid biosynthetic gene cluster on genomic DNA hybridizations to our microarrays. However, the observed overexpression of the clavams biosynthetic gene clusters—which are strongly related to the clavulanic acid gene cluster—in strain DS48802 may also point to recruitment of enzymes for clavulanic acid overproduction from these homologous pathways, alleviating the need for amplification of the original gene cluster. Even though in this case it would be an enigma why *cvm5* and *cvm6p*, which are thought to be specifically involved in the final steps of 5S clavam biosynthesis,[8,9] are also overexpressed in DS48802, the general picture shows that random

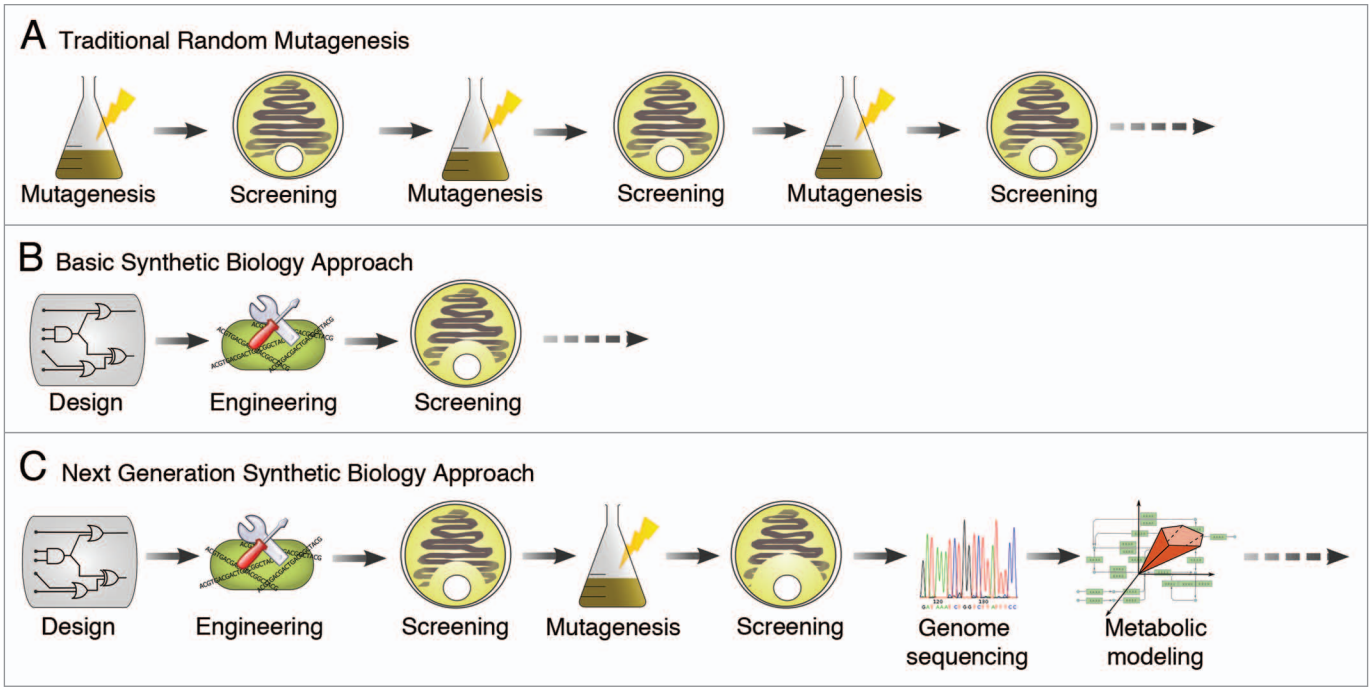

Figure 1. Comparison of several approaches for strain optimization. (A) In traditional random mutagenesis, the strain is subjected to several iterations of mutagenesis and screening, in an attempt to arrive at a higher-producing phenotype. (B) In the basic synthetic biology approach, an advantage is gained by intelligently engineering a high-producer strain. The strain is designed, engineered and an activity screen is performed to test the obtained production titers. (C) Proposed pipeline for compound production optimization based on a synergy of synthetic biology, systems biology and random mutagenesis. After initial engineering of a first design and screening of its productivity, a round of random mutagenesis of the resulting "designer bug version 1.0" follows. The mutants are screened and promising improved mutants are selected for genome sequencing. The identified mutations are tested, when possible, with genome-scale metabolic modeling and combined in engineering a second version of the designer bug. Several additional iterations of design updates or 'bug fixes', mutagenesis and genome sequencing may follow to further optimize the production titers.

mutagenesis and traditional engineering yield similar results: small adjustments to the metabolic or regulatory network of the cell which allow finding a local optimum of production given the rest of the network.

As the field of synthetic biology is maturing, the methodologies for rationally engineering bacterial strains for production of natural products are drastically changing.[10] With these changes come new prospects to make random mutagenesis approaches complement rational engineering in novel ways, in a 'next generation' synthetic biology approach. Synthetic biology engineering is by definition not restricted to the natural architecture of a certain bacterial strain: cellular systems can be extensively re-engineered, or even engineered from the ground up in a de novo fashion. But while the general biological knowledge necessary for major re-engineering is often available, the fine-tuning of a design up to the nucleotide level is often much more difficult. This is where random mutagenesis can come in, again to find a local optimum, but this time an optimum which may be much closer to the global optimum because the rest of the metabolic and regulatory network has first been radically modified.

As genome sequencing is becoming cheaper by the day, an interesting possibility would be to expose the first version of a design for an overproduction strain to one or a few rounds of mutagenesis, and subsequently sequence the genomes and transcriptomes of a range of promising mutants. In this way, mutations that can help optimize the first design may be identified and combined together without the need for long and tedious repetitions of such rounds of mutagenesis. Metabolic modeling can sometimes aid in the selection of the mutations from the modified genomes, by predicting whether the changes will lead to higher production of the compound. And conversely, the observed profile of mutations can pinpoint bottleneck reactions that may have been missed in the earlier model-driven engineering (**Fig. 1**).

When envisaging a Streptomyces host for synthetic biology, many efforts have been focussed on generating a minimal Streptomyces genome. For example, Komatsu et al.[11] recently engineered a genome-minimized genome of *Streptomyces avermitilis*, by deleting large regions from the chromosomal termini. Metabolic modeling may also aid in guiding the construction of such a minimal streptomycete, by predicting the minimal set of metabolic genes necessary for production of biomass. Recently, we used in silico knockouts in a *Streptomyces clavuligerus* genome-scale metabolic model to predict that the 1.8-Mb linear plasmid that the strain possesses is not required for primary metabolism and could potentially be cured from the strain.[12] We used the same model—the version published by Medema et al.[12] available from the authors in SBML format upon request—to

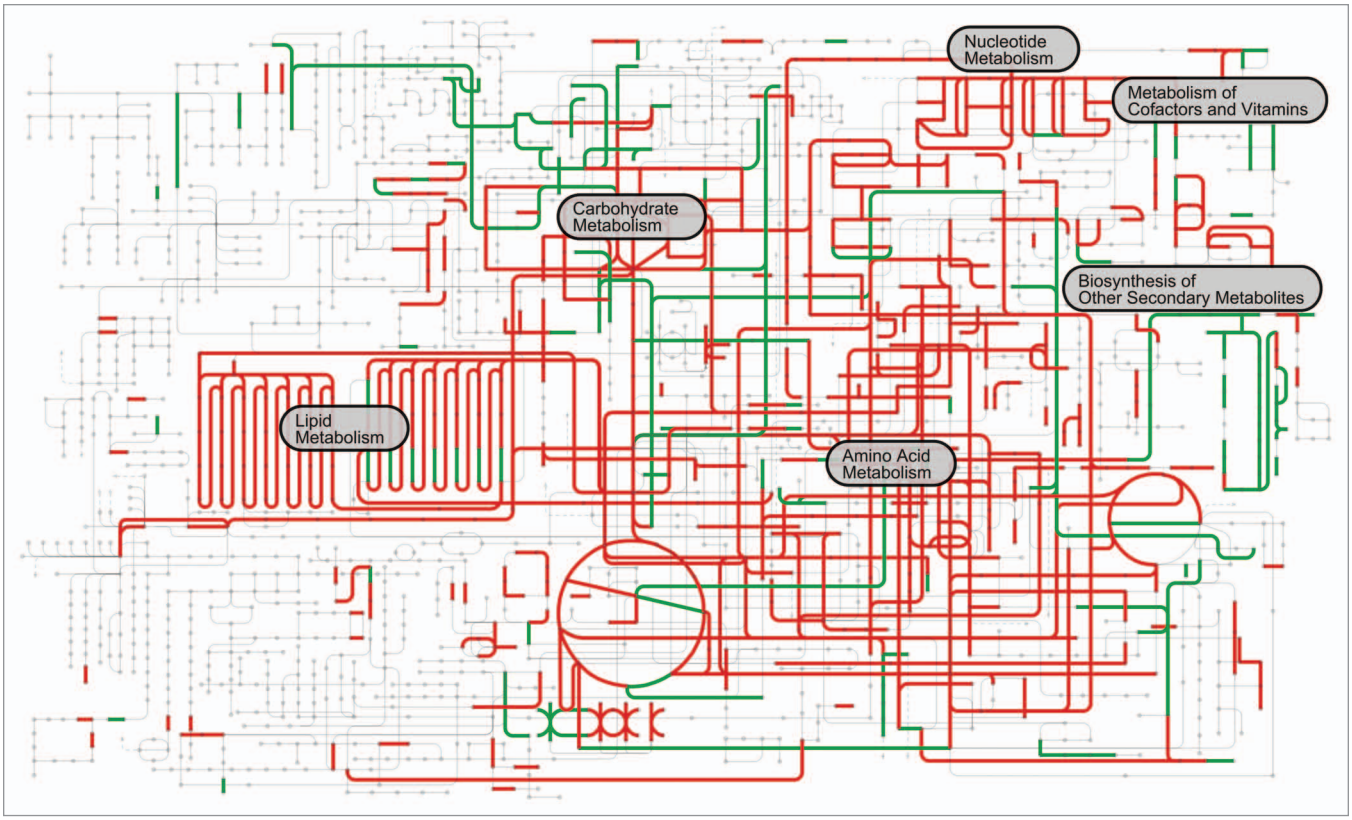

Figure 2. Visualization of essential and nonessential enzymes on the KEGG global pathway map, constructed using the iPATH 2 web server.[15] Reactions catalyzed by predicted essential enzymes are colored in red, reactions catalyzed by predicted non-essential enzymes are colored in green.

interpret the expression changes seen in the optimized strain[1], and again using flux balance analysis on in silico knockouts in a minimal medium, we have now attained a first approximation of the minimal *S. clavuligerus* genome. As a start, we generated in silico knockouts of all 507 unique enzymes (from 1115 EC-annotated enzymes in total) in the *S. clavuligerus* genome-scale metabolic model. For 159 knockouts (31%), the model was unable to simulate biomass production anymore; the 159 enzymes linked to these knockouts were therefore labeled as unconditionally essential. To be able to identify conditionally essential enzymes as well, we subsequently generated a metabolic network[13] which links enzyme nodes via edges based on shared common metabolites. Highly connected metabolites were removed from the model reactions, as well as reactions which contain one of these highly connected metabolites as their only product or substrate. We then performed double knockouts by removing all possible pairs of directly connected nodes (enzymes) from the network one by one, and checked the cellular growth again by performing flux balance analysis.[14] Of the 1,079 pairs of enzymes, 657 (61%) were predicted to be essential for biomass production. Seventy enzymes which were not essential according to the first analysis appeared to be essential when knocked out as part of a pair, and were therefore labeled as conditionally essential. Subsequently, we performed triple knockouts by removing three directly connected enzymes at a time. Nineteen new conditionally essential enzymes were detected from this round of knockouts. In total, 194 enzymes (38%) were found nonessential after these three rounds of in silico knockouts. However, after removing these 194 predicted nonessential enzymes from the genome-scale model, the model of the cell was not viable during flux balance analysis, probably because in a few cases the conditionality exists on an even higher level, e.g., when there exist two entire entirely independent alternative pathways of which at least one is necessary. Therefore, we manually gap-filled this reduced model by re-adding 49 enzymes, which connect essential enzymes with conditionally essential enzymes, from the set of 194 enzymes that were left from the previous analysis. This successfully allowed simulation of biomass production again. In the end, we identified a set of 145 enzymes (29%, **Sup. Table 1**) which were predicted to be nonessential. These enzymes are encoded by 195 chromosomally encoded genes and 20 plasmid-encoded genes which could potentially be deleted from the *S. clavuligerus* genome in order to minimize it. As the core chromosome of *S. clavuligerus* is already significantly smaller (6.8 Mb) than the recently published genome-minimized *S. avermitilis* chromosome (7.6 Mb),[11] the prospect of being able to further minimize the genome is enthralling. It should be kept in mind that these results were obtained from the analysis of only a fraction of the total 7,281 annotated genes from the *S. clavuligerus* genome (i.e., those that encode enzymes). Visualization of the predicted essential and nonessential

enzymes showed that the nonessential enzymes primarily reside in pathways involved in secondary metabolism, alternative carbon metabolisms and vitamin and cofactor biosynthesis. The latter could turn out to be essential on different minimal media. Potentially, this information could be used in an iterative strategy to cut away unnecessary chromosomal regions to further minimize the Streptomyces genome after the plasmid megaplasmid[12] could perhaps be cured out of the strain, or even—in the long run—to synthesize a Streptomyces genome de novo.

Overall, although classical strain improvement by random mutagenesis may be loosing its current key position in the generation of overproduction strains, the methodology may still remain valuable for fine-tuning the rough designs produced by synthetic biology. In silico modeling of metabolism is likely to complement these approaches by confirming the rationality of mutations, and providing new leads for altering the metabolic network, minimizing the genome and optimizing the whole system for overproduction of the compound of choice.

Note

Supplemental materials can be found at: www.landesbioscience.com/journals/biobugs/article/16114/

References

1. Medema MH, Alam MT, Heijne WH, van den Berg MA, Muller U, Trefzer A, et al. Genome-wide gene expression changes in an industrial clavulanic acid overproduction strain of *Streptomyces clavuligerus*. Microb Biotechnol 2011; 4:300-5.
2. Li R, Townsend CA. Rational strain improvement for enhanced clavulanic acid production by genetic engineering of the glycolytic pathway in *Streptomyces clavuligerus*. Metab Eng 2006; 8:240-52.
3. Hung TV, Malla S, Park BC, Liou K, Lee HC, Sohng JK. Enhancement of clavulanic acid by replicative and integrative expression of *ccaR* and *cas2* in *Streptomyces clavuligerus* NRRL3585. J Microbiol Biotechnol 2007; 17:1538-45.
4. Jnawali HN, Lee HC, Sohng JK. Enhancement of clavulanic acid production by expressing regulatory genes in *gap* gene deletion mutant of *Streptomyces clavuligerus* NRRL3585. J Microbiol Biotechnol 2010; 20:146-52.
5. Smanski MJ, Peterson RM, Rajski SR, Shen B. Engineered *Streptomyces platensis* strains that overproduce antibiotics platensimycin and platencin. Antimicrob Agents Chemother 2009; 53:1299-304.
6. Liao G, Li J, Li L, Yang H, Tian Y, Tan H. Cloning, reassembling and integration of the entire nikkomycin biosynthetic gene cluster into *Streptomyces ansochromogenes* lead to an improved nikkomycin production. Microb Cell Fact 2010; 9:6.

7. Yanai K, Murakami T, Bibb M. Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of *Streptomyces kanamyceticus*. Proc Natl Acad Sci USA 2006; 103:9661-6.
8. Tahlan K, Anders C, Wong A, Mosher RH, Beatty PH, Brumlik MJ, et al. 5S clavam biosynthetic genes are located in both the clavam and paralog gene clusters in *Streptomyces clavuligerus*. Chem Biol 2007; 14:131-42.
9. Zelyas NJ, Cai H, Kwong T, Jensen SE. Alanylclavam biosynthetic genes are clustered together with one group of clavulanic acid biosynthetic genes in *Streptomyces clavuligerus*. J Bacteriol 2008; 190:7957.
10. Medema MH, Breitling R, Bovenberg R, Takano E. Exploiting plug-and-play synthetic biology for drug discovery and production in microorganisms. Nat Rev Microbiol 2011; 9:131-7.
11. Komatsu M, Uchiyama T, Omura S, Cane DE, Ikeda H. Genome-minimized streptomyces host for the heterologous expression of secondary metabolism. Proc Natl Acad Sci USA 2010; 107:2646-51.
12. Medema MH, Trefzer A, Kovalchuk A, van den Berg M, Muller U, Heijne W, et al. The sequence of a 1.8-Mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. Genome Biol Evol 2010; 2:212-24.
13. Ma H, Zeng AP. Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. Bioinformatics 2003; 19:270-7.
14. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BO, Herrgard MJ. Quantitative prediction of cellular metabolism with constraint-based models: The COBRA toolbox. Nat Protoc 2007; 2:727-38.
15. Letunic I, Yamada T, Kanehisa M, Bork P. iPath: Interactive exploration of biochemical pathways and networks. Trends Biochem Sci 2008; 33:101-3.